# NOTES ON ESTIMATING VARIANCE COMPONENTS FROM UNBALANCED DATA IN MIXED MODELS OF THE ANALYSIS OF VARIANCE

BU-506-M

by

April, 1974

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

Abstract

An outline is given of 6 available methods for estimating variance components from unbalanced data in mixed models of the analysis of variance.

---

Paper No. BU-506-M in the Biometrics Unit, Cornell University, Ithaca, New York.

NOTES ON
ESTIMATING VARIANCE COMPONENTS FROM UNBALANCED DATA IN
MIXED MODELS OF THE ANALYSIS OF VARIANCE

BU-506-M

by

April, 1974

S. R. Searle

Biometrics Unit, Cornell University, Ithaca, New York

## Introduction

Confine attention to the 2-way crossed classification model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

$$i = 1 \cdots a \qquad j = 1 \cdots b \qquad k = 1 \cdots n_{ij} \qquad \sum_{ij}\sum n_{ij} = N$$

**Fixed effects model**

**Balanced data**

All $n_{ij} = n$: the analysis of variance is familiar.

| | | |
|---|---|---|
| Mean | 1 | $abn\bar{y}^2_{\cdot\cdot\cdot}$ |
| Rows | a-1 | $SSA = \Sigma bn\bar{y}^2_{i\cdot\cdot} - abn\bar{y}^2_{\cdot\cdot\cdot}$ |
| Columns | b-1 | $SSB = \Sigma an\bar{y}^2_{\cdot j\cdot} - abn\bar{y}^2_{\cdot\cdot\cdot}$ |
| Interaction | (a-1)(b-1) | $SSAB = \Sigma\Sigma n\bar{y}^2_{ij\cdot} - \Sigma bn\bar{y}^2_{i\cdot\cdot} - \Sigma an\bar{y}^2_{\cdot j\cdot} + abn\bar{y}^2_{\cdot\cdot\cdot}$ |
| Residual | ab(n-1) | $SSE = \Sigma\Sigma\Sigma y^2_{ijk} - \Sigma\Sigma n\bar{y}^2_{ij\cdot}$ |
| Total | abn | $\Sigma\Sigma\Sigma y^2_{ijk}$ |

**Unbalanced data**: s cells containing data.

2 partitionings of sums of squares.

| Rows before columns | | OR | Columns before rows | |
|---|---|---|---|---|
| $R(\mu)$ | 1 | | $R(\mu)$ | 1 |
| $R(\alpha\|\mu)$ | a-1 | | $R(\beta\|\mu)$ | b-1 |
| $R(\beta\|\mu,\alpha)$ | b-1 | | $R(\alpha\|\mu,\beta)$ | a-1 |
| $R(\gamma\|\mu,\alpha,\beta)$ | s-a-b+1 | | $R(\gamma\|\mu,\alpha,\beta)$ | s-a-b+1 |
| SSE | N-s | | SSE | N-s |
| Total | N | | Total | N |

Mixed model:

$\beta_j$'s remain as fixed effects

$\alpha_i$'s random

$E(\alpha_i) = 0$

$\mathrm{var}(\underset{\sim}{\alpha}) = \sigma^2_\alpha \underset{\sim}{I}_a$

$\gamma_{ij}$'s random

$E(\gamma_{ij}) = 0$

$\mathrm{var}(\underset{\sim}{\gamma}) = \sigma^2_\gamma \underset{\sim}{I}_s$

s = ab for balanced data

Want to estimate: $\mu$, $\beta$'s, $\sigma^2_\alpha$, $\sigma^2_\gamma$ and $\sigma^2_e$

Balanced data:

Use part of analysis of variance table for fixed effects model

$$E(SSA) = (a-1)(bn\sigma^2_\alpha + n\sigma^2_\gamma + \sigma^2_e)$$
$$E(SSAB) = (a-1)(b-1)(n\sigma^2_\gamma + \sigma^2_e)$$
$$E(SSE) = ab(n-1)\sigma^2_e$$

Estimators

$$SSA = (a-1)(bn\hat{\sigma}^2_\alpha + n\hat{\sigma}^2_\gamma + \hat{\sigma}^2_e)$$
$$SSAB = (a-1)(b-1)(n\hat{\sigma}^2_\gamma + \hat{\sigma}^2_e)$$
$$SSE = ab(n-1)\hat{\sigma}^2_e$$

Properties of estimators: unbiased

minimum variance quadratic unbiased

under normality, minimum variance unbiased

Unbalanced data:

Variety of methods available, several based on same principle as preceding:

Develop $\underset{\sim}{q}$ as a vector of quadratic forms in $\underset{\sim}{y}$

Derive $E(\underset{\sim}{q})$; each element will be a linear combination of variance components, elements of $\underset{\sim}{\sigma}^2$.

$$E(\underset{\sim}{q}) = \underset{\sim}{C}\underset{\sim}{\sigma}^2 \text{ for some } \underset{\sim}{C}$$

Estimation: $\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{C}^{-1}\underset{\sim}{q}$

Question: What quadratics are used as elements for $\underset{\sim}{q}$ ?

## 0. Analysis of variance method (Henderson's [1953] Method 1)

This method uses quadratic forms analogous to sums of squares of balanced data ANOVA, e.g.,

$$SSA^* = \sum_i n_{i.}\bar{y}^2_{i..} - N\bar{y}^2_{...}$$

$$SSAB^* = \sum\sum_{ij} n_{ij}\bar{y}^2_{ij.} - \sum_i n_{i.}\bar{y}^2_{i..} - \sum_j n_{.j}\bar{y}^2_{.j.} + N\bar{y}^2_{...}$$

Note: $SSAB^*$ is not positive definite; it is not a sum of squares.

Estimation: equate $SS^*$'s to expectations

Properties: easy to compute

unbiased for random models

sampling variances available for 1, 2 and 3-way classifications

not unbiased for mixed models, because the fixed effects, $\beta_j$'s, occur in $E(SS^*)$'s.

## 1. Henderson's [1953] Method 2

Designed to overcome biasedness of Method 1 for mixed models.

Retains relative ease of computing.

Principle: "Correct" data for fixed effects

Use Method 1 on corrected data

Make slight adjustments.

$$\underset{\sim}{y} = \underbrace{\underset{\sim}{X}\underset{\sim}{b}}_{\text{fixed}} + \underbrace{\underset{\sim}{Z}\underset{\sim}{u}}_{\text{random}} + \underset{\sim}{e}$$

Use normal equations as if $\underset{\sim}{u}$ were fixed:

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{X} & \underset{\sim}{Z}'\underset{\sim}{Z} \end{bmatrix} \begin{bmatrix} \underset{\sim}{b}^{\circ} \\ \underset{\sim}{u}^{\circ} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

Correct for $\underset{\sim}{b}$:

$$\underset{\sim}{z} = \underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}^{\circ} = \mu^{*}\underset{\sim}{1} + \underset{\sim}{Z}\underset{\sim}{u} + \underset{\sim}{K}\underset{\sim}{e}, \text{ for some } \underset{\sim}{K}.$$

Use Method 1 on $\underset{\sim}{z}$ just as if it were $\underset{\sim}{y}$ without fixed effects.

Adjustments: to coefficients of $\sigma_e^2$ in E(SS's), to account for $\underset{\sim}{K}$.

Condition: no interactions, fixed-by-random

History: Henderson [1953]: first described, and not clear.

Searle [1968]: generalized, clarified, decried as not invariant.

Henderson, Searle and Schaeffer [1974]: invariance established, and computing procedure described.

## 2. Fitting constants method (Henderson's [1953] Method 3)

Use R( )'s of fitting constants for fixed effects models

$$E\,R(\alpha,\gamma|\mu,\beta) = c_1\sigma^2_\alpha + c_1\sigma^2_\gamma + (s-b)\sigma^2_e$$

$$E\,R(\gamma|\mu,\alpha,\beta) = c_2\sigma^2_\gamma + (s-a-b+1)\sigma^2_e$$

$$E\,SSE = (N-s)\sigma^2_e$$

or, if no interaction

$$E\,R(\alpha|\mu,\beta) = c_1\sigma^2_\alpha + (a-1)\sigma^2_e$$

$$E\,SSE = (N-a-b+1)\sigma^2_e$$

Properties: unbiased

reduce to ANOVA for balanced

Difficulties: can be difficult to compute (i.e., inverting large matrices)

can have more equations than variance components

e.g., for 2-way random model, can use

| | | | | |
|---|---|---|---|---|
| $R(\alpha\|\mu)$ | | $R(\beta\|\mu)$ | | $R(\beta\|\mu,\alpha)$ |
| $R(\beta\|\mu,\alpha)$ | OR | $R(\alpha\|\mu,\beta)$ | OR | $R(\alpha\|\mu,\beta)$ |
| $R(\gamma\|\mu,\alpha,\beta)$ | | $R(\gamma\|\mu,\alpha,\beta)$ | | $R(\gamma\|\mu,\alpha,\beta)$ |
| SSE | | SSE | | SSE |
| $y'y - N\bar{y}^2$ | | $y'y - N\bar{y}^2$ | | |

As a preliminary to other methods consider the general model

$$\underset{\sim}{y} = \underbrace{\underset{\sim}{X}\underset{\sim}{b}}_{\text{fixed}} + \underbrace{\underset{\sim}{Z}\underset{\sim}{u}}_{\text{random}} + \underset{\sim}{e}$$

$$E(\underset{\sim}{u}) = \underset{\sim}{0} \qquad E(\underset{\sim}{e}) = \underset{\sim}{0}$$

$$\text{var}(\underset{\sim}{u}) = \underset{\sim}{D} \qquad \text{var}(\underset{\sim}{e}) = \underset{\sim}{R}$$

$$E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$$

$$\text{var}(\underset{\sim}{y}) = \underset{\sim}{Z}\underset{\sim}{D}\underset{\sim}{Z}' + \underset{\sim}{R} \equiv \underset{\sim}{V}$$

GLS for $\underset{\sim}{b}$:

$$\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X}\underset{\sim}{b}^{o} = \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

Difficulty: $\underset{\sim}{V}^{-1}$ of order N.

GLS for $\underset{\sim}{b}$ and $\underset{\sim}{u}$, assuming $\underset{\sim}{u}$ fixed

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\tilde{b}} \\ \underset{\sim}{\tilde{u}} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \end{bmatrix}$$

Amend equations by adding $\underset{\sim}{D}^{-1}$ to $\underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z}$:

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} \\ \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} & \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{Z} + \underset{\sim}{D}^{-1} \end{bmatrix} \begin{bmatrix} \underset{\sim}{b}^{*} \\ \underset{\sim}{u}^{*} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{R}^{-1}\underset{\sim}{y} \end{bmatrix}$$

These are sometimes called the "mixed model equations".

Can show that $\underset{\sim}{b}^{*}$ is the GLS of $\underset{\sim}{b}$: i.e., $\underset{\sim}{b}^{*} = \underset{\sim}{b}^{o}$

Special case: $\underset{\sim}{u} \equiv \underset{\sim}{\alpha}$, a single random factor, $\sigma^2_\alpha$, and $\text{var}(\underset{\sim}{e}) = \sigma^2_e\underset{\sim}{I}$

Define $\lambda = \sigma^2_e/\sigma^2_\alpha$ and $\underset{\sim}{P} = \underset{\sim}{Z}'\underset{\sim}{Z} + \lambda\underset{\sim}{I}$

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{X} & \underset{\sim}{X}'\underset{\sim}{Z} \\ \underset{\sim}{Z}'\underset{\sim}{X} & \underset{\sim}{P} \end{bmatrix} \begin{bmatrix} \underset{\sim}{b}^{*} \\ \underset{\sim}{u}^{*} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

## 3. Thompson's iterative method

Models with only 1 random factor; e.g., 2-way classification without interaction

$$y_{ijk} = \underbrace{\mu + \beta_j} + \alpha_i + e_{ijk}$$
$$\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{Z\alpha} + \underset{\sim}{e}$$

Fitting constants method: Based on

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z \end{bmatrix} \begin{bmatrix} b^o \\ \alpha^o \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad \text{and} \quad R(\mu,\alpha,\beta) = (b^{o\prime} \quad \alpha^{o\prime}) \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\hat{\sigma}^2_e = \frac{y'y - R(\mu,\alpha,\beta)}{N-a-b+1} \qquad \hat{\sigma}^2_\alpha = \frac{R(\alpha|\mu,\beta) - (a-1)\sigma^2_e}{N - \Sigma\Sigma n^2_{ij}/n_{i.}}$$

Cunningham and Henderson [1968]: Used mixed model equations

$$\begin{bmatrix} X'X & X'Z \\ Z'X & P \end{bmatrix} \begin{bmatrix} b^* \\ \alpha^* \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad \text{and} \quad R^*(\mu,\alpha,\beta) = (b^{*\prime} \quad u^{*\prime}) \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

and got

$$\overset{*}{\sigma}{}^2_e = \frac{y'y - R^*(\mu,\alpha,\beta)}{N-a-b+1} \qquad \overset{*}{\sigma}{}^2 = \frac{R^*(\mu,\alpha,\beta) - R(\mu,\beta) - (a-1)\overset{*}{\sigma}{}^2_e}{N + a\lambda - \Sigma\Sigma n^2_{ij}/n_{i.}}$$

Iterate on $\lambda = \sigma^2_e/\sigma^2_\alpha$ with $\underset{\sim}{P} \equiv \underset{\sim}{Z}'\underset{\sim}{Z} + \lambda\underset{\sim}{I}$ .

Thompson's [1969] method:

Located error in expectations of Cunningham and Henderson; correction yields

$$\tilde{\sigma}^2_e = \frac{y'y - R^*(\mu,\alpha,\beta)}{N - b} \qquad \tilde{\sigma}^2_\alpha = \frac{R^*(\mu,\alpha,\beta) - R(\mu,\beta)}{N - \Sigma\Sigma n^2_{ij}/n_{i.}}$$

Iterate on $\lambda = \sigma^2_e/\sigma^2_\alpha$ .

Computing formulae for 2-way, no interaction: Searle [1973].

Extension to 2-way, with interaction: Corbeil and Searle [1973].

(This is an extension from 1 to 2 random factors.)

## 4. MINQUE (Four papers by C. R. Rao)

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K+1} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta \qquad \theta = \text{factors A, B, } \cdots\text{, K, including interaction}$$

$$\underset{\sim}{u}_{K+1} \equiv \underset{\sim}{e}, \quad \underset{\sim}{Z}_{K+1} \equiv \underset{\sim}{I}_N$$

$$E(\underset{\sim}{u}_\theta) = \underset{\sim}{0}; \quad \text{var}(\underset{\sim}{u}_\theta) = \sigma^2_\theta \underset{\sim}{I}_{N_\theta}; \quad \text{and } \text{cov}(\underset{\sim}{u}_\theta \underset{\sim}{u}'_\varphi) = \underset{\sim}{0} \text{ for } \theta \neq \varphi .$$

Notation: $\underset{\sim}{V}_\theta = \underset{\sim}{Z}_\theta \underset{\sim}{Z}'_\theta$; $\underset{\sim}{W} = \sum_{\theta=A}^{K+1} \underset{\sim}{V}_\theta$; (Rao's $\underset{\sim}{V}$) $\underset{\sim}{V} = \sum_{\theta=A}^{K+1} \sigma^2_\theta \underset{\sim}{V}_\theta$ (Rao's $\underset{\sim}{V}^*$)

Estimation: by quadratics $\underset{\sim}{y}'\underset{\sim}{A}\underset{\sim}{y}$ with $\underset{\sim}{A}\underset{\sim}{X} = \underset{\sim}{0}$ choosing $\underset{\sim}{A}$:

MIVQUE: to minimize $2\text{tr}(VA)^2$ + term in $\underset{\sim}{A}$ and kurtosis parameters [1971a, p. 268; 1972, p. 113]

Under normality minimize $\text{tr}(VA)^2$: [1971b, pp. 447, 453].

$$\underset{\sim}{R} = \underset{\sim}{V}^{-1} - \underset{\sim}{V}^{-1}\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}$$

$$\underset{\sim}{S} = \{s_{\theta\varphi}\} = \{\text{tr}(\underset{\sim}{V}_\theta \underset{\sim}{R} \underset{\sim}{V}_\varphi \underset{\sim}{R})\} \text{ for } \theta,\varphi = A, B, \cdots, K+1$$

$$\underset{\sim}{u} = \{u_\theta\} = \{\underset{\sim}{y}'\underset{\sim}{R}\underset{\sim}{V}_\theta \underset{\sim}{R}'\underset{\sim}{y}\} \text{ for } \theta = A, B, \cdots, K+1$$

$$\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{S}^{-1}\underset{\sim}{u} .$$

Iterate on $\underset{\sim}{\sigma}^{2\prime} = (\sigma^2_A \ \sigma^2_B \cdots \sigma^2_K \ \sigma^2_e)$.

MINQUE: Use $\underset{\sim}{\sigma}^2 = \underset{\sim}{1}$ without iterating; i.e., use $\underset{\sim}{W}$ for $\underset{\sim}{V}$ in $\underset{\sim}{R}$. [1971a, p. 268; 1972, p. 113]

Example of use: Maddala and Mount [1973]

Generalizations: LaMotte [1973]

## 5. Maximum likelihood

Use normality and same model as MINQUE:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \sum_{\theta=A}^{K} \mathbf{Z}_\theta \mathbf{u}_\theta + \mathbf{e}$$

Notation:

$$\gamma_\theta = \sigma_\theta^2 / \sigma_e^2 \qquad\qquad \mathbf{H} = \mathbf{I}_N + \sum_{\theta=A}^{K} \gamma_\theta \mathbf{Z}_\theta \mathbf{Z}_\theta'$$

$$\operatorname{var}(\mathbf{y}) = \sigma_e^2 \mathbf{H} = \mathbf{V}.$$

Equations:

$$\mathbf{X}'\tilde{\mathbf{H}}^{-1}\mathbf{X}\tilde{\mathbf{b}} = \mathbf{X}'\tilde{\mathbf{H}}^{-1}\mathbf{y}$$

$$\tilde{\sigma}_e^2 = (\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}})'\mathbf{H}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}})/N$$

$$\operatorname{tr}(\tilde{\mathbf{H}}^{-1}\mathbf{Z}_\theta\mathbf{Z}_\theta') = (\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}})'\tilde{\mathbf{H}}^{-1}\mathbf{Z}_\theta\mathbf{Z}_\theta'\tilde{\mathbf{H}}^{-1}(\mathbf{y} - \mathbf{X}\tilde{\mathbf{b}})/\tilde{\sigma}_e^2, \quad \text{for } \theta = A, \ldots, K$$

For unbalanced data these equations have no solution; neither do they for some balanced data situations (e.g. 2-way crossed classification, random model, with interaction). Solutions must be confined to positive values.

History:

Hartley and Rao [1967]: Established equations, and solved (numerically) by steepest descent.
Hartley and Vaughn [1972]: Computer program, and small examples.
Harville [1975]: A comprehensive review.
Hemmerle and Hartley [1973]: Newton-Raphson, and a transformation.
Jennrich and Sampson [1976]: Discusses several algorithms.
Miller [1973]: Improved iterative procedure.

## 6. REML: Restricted Maximum Likelihood

Use normality and same model as ML and MINQUE:

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta + \underset{\sim}{e}$$

Define $\underset{\sim}{b}$ so that $\underset{\sim}{X}$ has full column rank. An easy definition is $\underset{\sim}{b} \equiv$ vector of population means of the filled sub-most cells of the fixed effects factors.

$k$ = number of filled cells, $n_t$ observations in $t^{th}$, $t = 1 \cdots k$

$\underset{\sim}{X} = \sum_{t=1}^{k}{}^{+} \underset{\sim}{1}_{n_t}$, a direct (Kronecker) sum of $\underset{\sim}{1}$-vectors

$$\underset{\sim}{S} = \underset{\sim}{I} - \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}' = \sum_{t=1}^{k}{}^{+} (\underset{\sim}{I}_{n_t} - \frac{1}{n_t}\underset{\sim}{J}_{n_t})$$

$\underset{\sim}{T} = \underset{\sim}{S}$ after deleting rows $n_1, (n_1 + n_2), \ldots, (n_1 + n_2 + \cdots + n_k)$.

$$\underset{\sim}{z} = \begin{bmatrix} \underset{\sim}{T}\underset{\sim}{y} \\ \underset{\sim}{X}'\underset{\sim}{H}^{-1}\underset{\sim}{y} \end{bmatrix} \sim N\left[\begin{pmatrix} \underset{\sim}{0} \\ \underset{\sim}{X}'\underset{\sim}{H}^{-1}\underset{\sim}{X}\underset{\sim}{b} \end{pmatrix}, \begin{pmatrix} \underset{\sim}{T}\underset{\sim}{H}\underset{\sim}{T}'\sigma_e^2 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{X}'\underset{\sim}{H}^{-1}\underset{\sim}{X}\sigma_e^2 \end{pmatrix}\right]$$

To estimate $\underset{\sim}{\sigma}^2$, maximize the likelihood of $\underset{\sim}{T}\underset{\sim}{y}$, which does not involve $\underset{\sim}{b}$.

History:

Patterson and Thompson [1971]: Initial ideas, confined to b.i.b. designs.
Hocking and Kutner [1975]: Simulations on a b.i.b. design.
Harville [1975]: Comprehensive review.
Corbeil and Searle [1976]: Generalization, and computing procedures using Hemmerle and Hartley [1973].
Corbeil and Searle [1977]: Analytic comparisons for balanced data and numeric comparisons for unbalanced data.

## 7. Relationships among Methods

(1) ANOVA ≡ Henderson 1 (Definition).

(2) ML estimators are ML solutions subject to non-negativity conditions.

### Balanced Data

(3) ANOVA = Henderson 2 = Henderson 3 = REML = MINQUE (MIVQUE).

(4) Some ML equations have no closed form solution. When solutions do exist, some (but not all) = ANOVA. (Differences occur in "degrees of freedom").

### Unbalanced Data

ML and MINQUE (MIVQUE) model:

$$\underset{\sim}{y}_{N\times 1} = \underset{\sim}{X}\underset{\sim}{b} + \sum_{\theta=A}^{K} \underset{\sim}{Z}_\theta \underset{\sim}{u}_\theta + \underset{\sim}{e}, \qquad \text{with } \underset{\sim}{u}_\theta \text{ order } n_\theta \times 1.$$

Henderson's mixed model equations (HMME's)

$$\gamma_\theta = \sigma_\theta^2 / \sigma_e^2 \qquad \underset{\sim}{D} = \operatorname{diag}\{\gamma_\theta \underset{\sim}{I}_{n_\theta}\} \qquad \theta = A,\ldots K$$

$$\begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{X} & X'Z \\ \underset{\sim}{Z}'\underset{\sim}{X} & \underset{\sim}{Z}'\underset{\sim}{Z} + \underset{\sim}{D}^{-1} \end{bmatrix} \begin{bmatrix} b^* \\ u^* \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{Y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

with solution

$$\begin{bmatrix} \underset{\sim}{b}^* \\ \underset{\sim}{u}^* \end{bmatrix} = \begin{bmatrix} \underset{\sim}{C}_{00} & \underset{\sim}{C}_{01} \\ \underset{\sim}{C}_{10} & \underset{\sim}{C}_{11} \end{bmatrix} \begin{bmatrix} \underset{\sim}{X}'\underset{\sim}{y} \\ \underset{\sim}{Z}'\underset{\sim}{y} \end{bmatrix}$$

where

$$\underset{\sim}{C}_{11} = \{\underset{\sim}{C}_{\theta\phi}\} \qquad \theta,\phi = A,\ldots,K$$

$$\underset{\sim}{u}^* = \{\underset{\sim}{u}_\theta\} \qquad \theta = A,\ldots,K$$

MINQUE (MIVQUE) and the HMME's

MINQUE equations

$$\underset{\sim}{S}\hat{\underset{\sim}{\sigma}}^2 = \underset{\sim}{q}, \quad \text{i.e.} \quad \{s_{\theta\phi}\}\hat{\underset{\sim}{\sigma}}^2 = \{q_\theta\} \text{ for } \theta,\phi = A,\ldots,K,e$$

are given by

$$(5) \quad s_{\theta\theta} = [\mathrm{tr}(\underset{\sim}{C}_{\theta\theta})/\gamma_\theta^2 - 2\mathrm{tr}(\underset{\sim}{C}_{\theta\theta})/\gamma_\theta + n_\theta]/\gamma_\theta^2$$

$$s_{\theta\phi} = \mathrm{tr}(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\theta^2\gamma_\phi^2$$

$$s_{\theta e} = s_{e\theta} = \mathrm{tr}[\underset{\sim}{C}_{\theta\theta} - \sum_{\phi=A}^{K}\mathrm{tr}(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\phi]/\gamma_\theta$$

$$s_{ee} = N - r(\text{HMME's}) + \sum_{\theta=A}^{K}\sum_{\phi=A}^{K}\mathrm{tr}(\underset{\sim}{C}_{\theta\phi}\underset{\sim}{C}_{\phi\theta})/\gamma_\theta\gamma_\phi$$

$$q_\theta = \underset{\sim}{u}_\theta^{*\prime}\underset{\sim}{u}_\theta^{*}/\gamma_\theta^2$$

$$q_e = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{b}^{*\prime}\underset{\sim}{X}'\underset{\sim}{y} - \underset{\sim}{u}^{*\prime}\underset{\sim}{Z}'\underset{\sim}{y} - \sum_{\theta=A}^{K}\gamma_\theta q_\theta$$

ML and the HMME's

(6) Iterate ML using

$$\tilde{\sigma}_e^2 = \frac{\underset{\sim}{y}'(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}^* - \underset{\sim}{Z}\underset{\sim}{u}^*)}{N} \quad \text{and} \quad \tilde{\sigma}_\theta^2 = \frac{\underset{\sim}{u}_\theta^{*\prime}\,\underset{\sim}{u}_\theta^{*}}{n_\theta - \mathrm{tr}(\underset{\sim}{T}_{\theta\theta})}$$

for $(\underset{\sim}{I} + \underset{\sim}{Z}'\underset{\sim}{Z}\underset{\sim}{D})^{-1} = \{\underset{\sim}{T}_{\theta\phi}\}, \quad \theta,\phi = A,\ldots,K.$

This iteration always gives positive estimators.

REML and MINQUE (MIVQUE)

(7) REML equations = MINQUE equations.

(8) REML estimators = Iterative MINQUE estimators.

(9) First iterate of REML = A MINQUE estimator.

History:

Patterson and Thompson [1971]: First indication of result for $q_\theta$, for b.i.b. design.
Henderson [1973]: Extended results, HMME's, ML and MINQUE.
La Motte [1973]: Indicated results for REML and MINQUE.
Schaeffer [1975]: Published same details HMME's and MINQUE.
Harville [1975]: Comprehensive review.

## REFERENCES


0. Analysis of variance method (Henderson's Method 1)

(For random models but not mixed models)

Henderson, C. R. [1953]. Estimation of variance and covariance components. Biometrics 9, 226-52.

Searle, S. R. [1968]. Another look at Henderson's methods of estimating variance components. Biometrics 24, 749-88.

Searle, S. R. [1971a]. Linear Models, Wiley.

Searle, S. R. [1971b]. Topics in variance components estimation. Biometrics 27, 1-77.

1. Henderson's Method 2 (for models with no fixed-by-random interactions)

Henderson, C. R. [1953].

Searle, S. R. [1968, 1971a, 1971b].

Henderson, C. R., S. R. Searle and L. R. Schaeffer [1974]. The invariance and calculation of Method 2 for estimating variance components. Biometrics 30, 583-588.

2. Fitting constants method (Henderson's Method 3)

Henderson, C. R. [1953].

Searle, S. R. [1968, 1971a, 1971b].

3. Thompson's iterative method

Cunningham, E. P. and Henderson, C. R. [1968]. An iterative procedure for estimating fixed effects and variance components in mixed model situations. Biometrics 24, 13-25. Correction 25, 777-78.

Thompson, R. [1969]. Iterative estimation of variance components for non-orthogonal data. Biometrics 25, 767-73.

Searle, S. R. [1971a].

Searle, S. R. [1973]. Computing procedures for estimating variance components from unbalanced data in the 2-way crossed classification, no interaction, mixed model. Paper BU-450-M in the Biometrics Unit, Cornell University.

Corbeil, R. R. and Searle, S. R. [1973]. Iterative estimation of variance components in the 2-way crossed classification mixed model, with interaction, using unbalanced data. Paper BU-460-M in the Biometrics Unit, Cornell University.

4. MINQUE

LaMotte, L. R. [1973]. Quadratic estimation of variance components. Biometrics 29, 311-330.

Maddala, G. S. and Mount, T. D. [1973]. Comparative study of alternative estimators for variance components models. J.A.S.A. 68, 324-328.

Rao, C. R. [1970]. Estimation of heteroscedastic variances in linear models. J. Am. Stat. Assoc. 65, 161-172.

Rao, C. R. [1971a]. Estimation of variance and covariance components - MINQUE theory. J. Multivar. Anal. 1, 257-275.

Rao, C. R. [1971b]. Minimum variance quadratic unbiased estimation of variance components. J. Multivar. Anal. 1, 445-456.

Rao, C. R. [1972]. Estimation of variance and covariance components in linear models. J. Am. Stat. Assoc. 67, 112-115.

5. Maximum Likelihood

Hartley, H. O. and Rao, J. N. K. [1967]. Maximum likelihood estimation for the mixed analysis of variance model. Biometrika 54, 93-108.

Hartley, H. O. and Vaughn, W. K. [1972]. A computer program for the mixed analysis of variance model based on maximum likelihood. In Statistical Papers in Honor of George W. Snedecor, Ed., T. A. Bancroft, 129-144, Iowa State University Press, Ames.

Harville, D. A. [1975]. Maximum likelihood approaches to variance component estimation and related problems. Technical Report No. 75-0175, Aerospace Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Hemmerle, W. J. and Hartley, H. O. [1973]. Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. Technometrics 15, 819-832.

Jennrich, R. I. and Sampson, P. F. [1976]. Newton-Raphson and related algorithms for maximum likelihood estimation of variance components. Technometrics, 18, 11-18.

Miller, J. J. [1973]. Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance. Tech. Rpt. #12, (NR-042-034), Department of Statistics, Stanford University, Stanford, California.

## 6. Restricted Maximum Likelihood

Corbeil, R. R. and Searle, S. R. [1976]. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. Technometrics 18, 31-38.

Corbeil, R. R. and Searle, S. R. [1977]. A comparison of variance components estimators. Biometrics, (in press).

Harville, D. A. [1975].

Patterson, H. D. and Thompson, R. [1971]. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554.

## 7. Relationships among ML, REML and MINQUE

Harville, D. A. [1975].

Henderson, C. R. [1973]. (i) MINQUE of variance components; and (ii) Maximum likelihood estimation of variance components. Mimeo, Animal Science Department, Cornell University.

LaMotte [1973]. Quadratic estimation of variance components. Biometrics 29, 311-330.

Schaeffer, L. R. [1975]. Disconnectedness and variance component estimation. Biometrics 31, 969-977.